OXFORD

BJS, 2021, **108**, 315–325
DOI: 10.1093/bjs/znaa125
Advance Access Publication Date: 24 March 2021
Original Article

Bridging the age gap in breast cancer: impact of omission of breast cancer surgery in older women with oestrogen receptor-positive early breast cancer on quality-of-life outcomes

J. L. Morgan [1], A. Shrestha[1], M. W. R. Reed [2], E. Herbert[3], M. Bradburn[3], S. J. Walters[3], C. Martin[1], K. Collins[4], S. Ward [5], G. Holmes [5], M. Burton[3], K. Lifford[6], A. Edwards[6], A. Ring[7], T. Robinson[8], T. Chater[3], K. Pemberton[3], A. Brennan[5], K. L. Cheung[9], A. Todd[1], R. Audisio[10], J. Wright[2], R. Simcock[2], A. M. Thomson[11], M. Gosney[12], M. Hatton[13], T. Green[14], D. Revill[14], J. Gath[14], K. Horgan[15], C. Holcombe[16], M. C. Winter[13], J. Naik[17], R. Parmeschwar[18] and L. Wyld [1,*]

[1]Department of Oncology and Metabolism, University of Sheffield Medical School, Sheffield, UK
[2]Brighton and Sussex Medical School, Brighton, UK
[3]Clinical Trials Research Unit, School for Health and Related Research, University of Sheffield, Sheffield, UK
[4]Faculty of Health and Wellbeing, Department of Allied Health Professions, Sheffield Hallam University, Sheffield, UK
[5]Department of Health Economics and Decision Science, School for Health and Related Research, ScHARR, University of Sheffield, Sheffield, UK
[6]Division of Population Medicine, Cardiff University, Cardiff, UK
[7]Department of Medical Oncology, Royal Marsden Hospital, London, UK
[8]Department of Cardiovascular Sciences and National Institute for Health Research Biomedical Research Centre, University of Leicester, Cardiovascular Research Centre, Glenfield General Hospital, Leicester, UK
[9]School of Medicine, University of Nottingham, Royal Derby Hospital, Derby, UK
[10]Department of Surgery, University of Gothenberg, Sahlgrenska Universitetssjukhuset, Gothenberg, Sweden
[11]Department of Surgery, Baylor College of Medicine, Houston, Texas, USA
[12]School of Psychology and Clinical Language Sciences, University of Reading, Reading, UK
[13]Sheffield Teaching Hospitals NHS Foundation Trust, Weston Park Hospital, Sheffield, UK
[14]North Trent Cancer Research Network Consumer Research Panel, Sheffield, UK
[15]Department of Breast Surgery, Bexley Cancer Centre, St James's University Hospital, Leeds, UK
[16]Department of Breast Surgery, Liverpool University Hospitals Foundation Trust, Liverpool, UK
[17]Department of General Surgery, Pinderfields Hospital, Mid Yorkshire NHS Foundation Trust, Wakefield, UK
[18]Department of Breast Surgery, University Hospitals of Morecambe Bay, Lancaster, UK

Correspondence to: Department of Oncology and Metabolism, University of Sheffield Medical School, Beech Hill Road, Sheffield S10 2RX, UK (e-mail: l.wyld@sheffield.ac.uk)

Abstract

Background: Primary endocrine therapy may be an alternative treatment for less fit women with oestrogen receptor (ER)-positive breast cancer. This study compared quality-of-life (QoL) outcomes in older women treated with surgery or primary endocrine therapy.

Methods: This was a multicentre, prospective, observational cohort study of surgery or primary endocrine therapy in women aged over 70 years with operable breast cancer. QoL was assessed using European Organisation for Research and Treatment of cancer QoL questionnaires QLQ-C30, -BR23, and -ELD14, and the EuroQol Five Dimensions 5L score at baseline, 6 weeks, and 6, 12, 18, and 24 months. Propensity score matching was used to adjust for baseline variation in health, fitness, and tumour stage.

Results: The study recruited 3416 women (median age 77 (range 69–102) years) from 56 breast units. Of these, 2979 (87.2 per cent) had ER-positive breast cancer; 2354 women had surgery and 500 received primary endocrine therapy (125 were excluded from analysis due to inadequate data or non-standard therapy). Median follow-up was 52 months. The primary endocrine therapy group was older and less fit. Baseline QoL differed between the groups; the mean(s.d.) QLQ-C30 global health status score was 66.2(21.1) in patients who received primary endocrine therapy *versus* 77.1(17.8) among those who had surgery plus endocrine therapy. In the unmatched analysis, changes in QoL between 6 weeks and baseline were noted in several domains, but by 24 months most scores had returned to baseline levels. In the matched analysis, major surgery (mastectomy or axillary clearance) had a more pronounced adverse impact than primary endocrine therapy in several domains.

Conclusion: Adverse effects on QoL are seen in the first few months after surgery, but by 24 months these have largely resolved. Women considering surgery should be informed of these effects.

Received: July 14, 2020. **Accepted:** November 15, 2020

Introduction

Although surgery is a key element of treatment for early breast cancer, it may sometimes be avoided in the frail older patient, who may instead be managed with primary endocrine therapy. Use of primary endocrine therapy instead of surgery was first suggested by Preece and colleagues[1] in 1980, and has been adopted variably across the globe[2]; some countries treat up to 40 per cent of women aged over 70 years using primary endocrine therapy, whereas others, such as the USA, use it rarely. The rationale is that in oestrogen receptor (ER)-positive cancer, disease control may be achieved for several years with primary endocrine therapy, with minimal morbidity. Frail older women with a shortened life expectancy may die from other causes before the cancer develops resistance to endocrine therapy, and so any adverse impact of surgery on quality of life (QoL) can be avoided. In particular, a frail older patient undergoing general anaesthesia may suffer a long-term reduction in functional capacity, along with risks of local and systemic complications. Numerous studies have shown that older patients value QoL at least as highly as length of life, and wish to maintain their QoL and independence for as long as possible to avoid being a burden to family and carers[3].

The evidence base for the health threshold at which primary endocrine therapy becomes appropriate is limited. European Society of Breast Cancer Specialists and the Society of Geriatric Oncology guidelines[4] recommend surgery for all women who are fit enough, but do not specify any health or fitness parameters for this assessment.

There are few published data on the QoL impact of primary endocrine therapy *versus* surgery, The Cancer Research Campaign trial[5], one of the early randomized trials comparing surgery and primary endocrine therapy, included a QoL substudy that used the General Health Questionnaire score up to 2 years after diagnosis. The median interval until first assessment was 12 months after treatment had started in both arms, and therefore at a time when much of the impact of surgery would have resolved. The study found that 9 of 49 patients in the surgery group and 2 of 49 in the primary endocrine therapy group had scores indicative of psychiatric morbidity; by 2 years, the numbers were 6 of 49 in both arms. Moreover, at the time of the Cancer Research Campaign trial, there were no validated tools for assessment of the specific impact of breast cancer and its treatment nor questionnaires evaluating more general issues, such as ability to function normally. Subsequently, the European Organisation for the Research and Treatment of Cancer (EORTC) developed a generic cancer QoL instrument (QLQ-C30)[6], and a breast cancer-specific instrument (QLQ-BR23)[7] with domains specific to breast cancer treatment effects. Another EORTC instrument has been developed more recently specifically to record QoL concerns relevant to elderly patients (QLQ-ELD15)[8]. No studies have compared QoL outcomes between surgery and primary endocrine therapy using these highly specific and validated tools.

A previous study[3] of older women who had faced the choice of primary endocrine therapy *versus* surgery demonstrated that disability-free life expectancy, QoL, and independence were key priorities[3], along with fear of surgery, anaesthesia, and hospitalization. Although surgery and anaesthesia have become safer in recent years, and breast surgery in particular is associated with a very low mortality rate[9], morbidity rates may be substantial[10]. A recent study[11] from the USA of frail older women with breast cancer reported that those living in nursing homes had high rates of mortality from surgery, significant rates of death within 1 year (implying that the surgery was of negligible benefit), and significant rates of functional decline after surgery. In this frailer older group surgery may therefore be harmful and convey little benefit, but impair highly valued QoL.

The Age Gap study was designed to determine age and health-stratified outcomes (including QoL) of surgery or primary endocrine therapy in older women with early breast cancer. The aim of the present study was to use a range of validated tools to provide a comprehensive evaluation of the impact of treatment, specifically comparing surgery plus endocrine therapy *versus* primary endocrine therapy, for early ER-positive breast cancer in older women, with a focus on factors including disability-free life expectancy, QoL, and independence.

Methods

Ethics and research governance approvals were obtained (IRAS: 12 LO 1808). All patients gave written informed consent to participate. The study was sponsored by Doncaster and Bassetlaw Teaching Hospitals NHS Foundation Trust, and registered as ISRCTN46099296.

Study design

This was a prospective, longitudinal, multicentre observational cohort study. Patients could participate at three levels: full participation, partial (no requirement to complete QoL assessments) or by proxy consent (simple third-party data collection for women with cognitive impairment). Trial reporting complies with STROBE guidelines[12]. The data in this paper relate only to women who consented to full participation.

Patients

Patients were recruited from 56 breast units in England and Wales (*Table S1*). Inclusion criteria were: women aged 70 years age or above at the time of breast cancer diagnosis with primary operable invasive breast cancer (TNM stages T1–3 and some T4b N0–2 M0). Women with multifocal and bilateral cancers were eligible. Exclusion criteria were: inoperable disease, and previous breast cancer within 5 years.

Baseline health assessment

Women were recruited at the time of breast cancer diagnosis and before commencement of treatment. Baseline health assessment was undertaken before treatment started using validated tools to permit baseline adjustment and reduce allocation bias. This included evaluation of co-morbidities (Charlson co-morbidity index, CCI[13]); nutrition (Abridged Patient Generated Subjective Global Assessment, aPG-SGA[14–16]); physical function (Barthel Activities of Daily Living, ADL)[17]; complex physical functioning (Instrumental Activities of Daily Living, IADL[18]; cognitive status (Mini-Mental State Examination, MMSE[19]); Eastern Cooperative Oncology Group Performance status, ECOG-PS)[20]; medication; and QoL assessment (see below).

Quality-of-life assessment

Four validated instruments were used for QoL assessment (*Table S2*). The EORTC QLQ-C30[6] generic QoL instrument has 28 questions covering seven domains, including physical, emotional, and role function, and two visual analogue scales (VAS) rating health status and QoL today. The EORTC-QLQ-BR23[7], is a breast-specific QoL module, which has 23 questions covering domains relevant to breast cancer therapies such as body image, breast symptoms, arm symptoms, chemotherapy, and endocrine

therapy symptoms. The EORTC-QLQ-ELD15[8] is an older person-specific module, which has 15 questions relating to five domains including mobility, and maintaining autonomy and purpose. Finally, the EuroQol Five Dimensions 5L (EQ-5D-5L™; EuroQoL Group, Rotterdam, the Netherlands)[21] is a generic instrument, with five domains including mobility, self-care and usual activities, and a VAS. (It was also used for health economic evaluation, which is reported separately). Measures of independence were taken from the usual activities section of the EQ-5D™ and some of the functional domains of the other questions.

Tumour and treatment data

The following baseline tumour data were collected: cancer type, grade, nodal status, tumour size (clinical and on imaging), oestrogen, progesterone, and human epidermal growth factor receptor 2 status, and Oncotype DX scores (if available). Staging for metastatic disease was performed if indicated clinically, but otherwise M0 was presumed.

Type of surgery, use of radiotherapy, and use and type of any systemic therapy were recorded. Baseline assessment of the tumour was undertaken using the Response Evaluation Criteria in Solid Tumours[22] to permit monitoring of the response to primary endocrine therapy. Type of surgery was classified as major (mastectomy and/or axillary clearance) or minor (wide local excision and/or sentinel lymph node biopsy) for some evaluations as a *post hoc* analysis.

Follow-up

Patients were followed up at 1.5, 6, 12, 18, and 24 months, with QoL assessment at each of these time points

Statistical analysis

The QLQ-C30, -BR23, and -ELD15 were scored according to the EORTC Scoring Manual (3rd edition)[23] and reference publications[8,24]. The EQ-5D-5L™ was used as a QoL measure for this analysis, with individual questions scored from 1 to 5; higher scores indicated better responses. An overall score from 0 to 1 was calculated, where higher scores indicated better QoL[25]. Missing data were managed according to EORTC recommendations.

$P < 0.050$ was was considered statistically significant. Analyses were done in SPSS® version 24 (IBM, Armonk, New York, USA), R version 3.6.3 (R Foundation for Statistical Computing, Vienna, Austria) and Stata® (StataCorp, College Station, Texas, USA).

Propensity score matching

Propensity score matching was used to identify patients with similar baseline characteristics in groups of patients undergoing surgery plus endocrine therapy or primary endocrine therapy, in an attempt to correct for allocation bias and allow a more direct comparison between groups.

Propensity scores for treatment allocation were calculated by means of logistic regression, and then used to match patients in the two treatment groups. The co-variables included were measures of functionality (ADL, IADL, MMSE, ECOG-PS), nutritional status (nutrition (aPG-SGA), co-morbidities (CCI, number of medications), and age (*Table S3*).

The ratio and caliper widths of the propensity scores were chosen based on examination of propensity score overlaps for several combinations of ratios and calipers. A 1 : 2 ratio for primary endocrine therapy to surgery plus endocrine therapy, and a caliper of 0.25 times the propensity scores standard deviation were used to optimally match quality and numbers. Participants were also matched exactly with respect to Nottingham Prognostic Index[26] category (good, below 3.5; moderate, 3.5–5.4; poor, over 5.4) to avoid fit participants with aggressive cancer being matched with frail participants who had a smaller cancer.

The impact of major or minor surgery *versus* primary endocrine therapy on domains of interest from each questionnaire was compared by fitting longitudinal mixed-effect models to the propensity score-matched data with treatment (minor surgery, major surgery or primary endocrine therapy), time, treatment–time interaction, and the baseline score as co-variables. The models were then used to estimate marginal effects of minor or major surgery *versus* primary endocrine therapy at each time point.

Results

The study recruited 3416 women between February 2013 and June 2018, with a median age of 77 (range 69–102) years (*Fig. 1*). Of these, only 3315 were fully eligible for analysis. The majority (2854, 86.1 per cent) had ER-positive breast cancer and were therefore potential candidates for the analysis of surgery plus endocrine therapy *versus* primary endocrine therapy. The following results relate only to patients with ER-positive breast cancer. Survival and recurrence outcomes have been published separately[10]. Broadly, in the unmatched unadjusted analyses, overall survival (OS) and breast cancer-specific survival (BCSS) were poorer in the primary endocrine therapy group than among patients who had surgery and endocrine therapy (OS: HR 0.27, 95 per cent c.i. 0.23 to 0.33, $P < 0.001$; BCSS: HR 0.89, 0.52 to 1.53, $P = 0.68$). On matching, these differences decreased (OS: adjusted HR 0.72, 0.53 to 0.98; $P = 0.037$) or disappeared (BCSS: HR 0.74, 0.40 to 1.37; $P = 0.340$), suggesting that, in a cohort of older frailer women, there was no survival advantage to surgery at 52 months' follow-up.

Of the 2854 patients with ER-positive breast cancer, 2354 underwent surgery (60.1 per cent breast conservation, 39.9 per cent mastectomy[10]) and 500 commenced primary endocrine therapy (82.8 per cent with letrozole, 6.2 per cent with anastrazole, 4.4 per cent with tamoxifen, 1.4 per cent with exemestane, 5.2 per cent type unknown). Of the 2354 surgical patients, 255 (10.8 per cent) also had chemotherapy within 12 months of diagnosis, and 1363 (57.9 per cent) underwent radiotherapy within 12 months of diagnosis. The impact of adjuvant chemotherapy and radiotherapy on QoL has been analysed and will be published separately. Patient, tumour, and treatment characteristics are summarized in *Tables 1* and *2*. Patients undergoing surgery were generally younger, fitter, and had better functionality than those undergoing primary endocrine therapy. Women in the surgery group had additional therapies including endocrine therapy, radiotherapy, chemotherapy, and trastuzumab variously according to clinical indication at various time points (*Fig. S1*).

Most surgery had taken place by the 6-week time point, although a small number of primary operations were undertaken between 6 weeks and 6 months (*Fig. S1*). For the purpose of this analysis, therefore, the impact of surgery is presumed to be represented by the change between baseline and 6 weeks.

Matching was undertaken to reduce baseline variation in health status, resulting in a matched population of 238 patients who had primary endocrine therapy (47.6 per cent), 184 of whom were matched with two surgical patients and 54 with only one. Matching quality was good (*Table S3*).

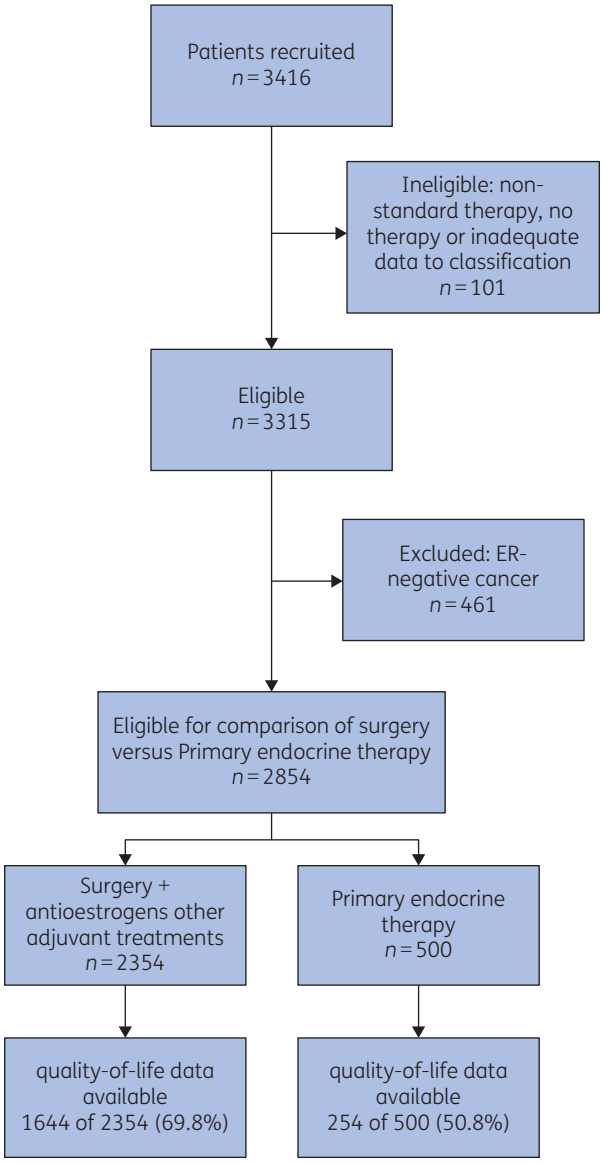

Fig. 1 Study flow diagram

Data completeness

QoL data were available for 1644 of 2354 surgical patients (69.8 per cent) and 254 of 500 (50.8 per cent) in the primary endocrine therapy group. Not all women completed 2-year follow-up because of discontinuation (deaths, withdrawals). Finally, some forms were not completed or completed inadequately, meaning that full data were not available for all patients. Owing to this attrition in numbers completing forms, especially in the primary endocrine therapy group, statistical interpretation was not possible for some domains; however, clear and clinically relevant variations and trends were seen in the data, as presented below.

Because of the variation in baseline QoL in all domains (owing to the different characteristics of women in each group), unmatched data are presented so that temporal trends resulting from treatment can be seen in all-comers (the matched group, plus very fit surgical patients and very unfit patients who received primary endocrine therapy for whom matching was not possible). Statistical comparison of these raw unmatched data was not possible owing to the wide variation in baseline QoL; therefore, statistical analysis was performed only on the matched data set in which baseline variation was minimized, making direct comparison of surgery plus endocrine therapy *versus* primary endocrine therapy more appropriate.

Generic cancer quality-of-life outcomes: EORTC-QLQ-C30

The proportion of patients with full participation who completed the QLQ-C30 global health status score at baseline was 1902 of 2065 (92.1 per cent), of whom 1644 of 1772 (92.8 per cent) had surgery plus endocrine therapy and 258 of 293 (88.1 per cent) primary endocrine therapy. At 24 months, these numbers had reduced to 984 of 2065 (47.7 per cent) in total, of whom 902 of 1772 (50.9 per cent) had surgery plus endocrine therapy and 82 of 293 (28.0 per cent) primary endocrine therapy. Full data (mean(s.d.) scores at all time points with numbers of completions) are shown in *Table S4*. Considering the differences in age, and health and fitness characteristics beween women treated with primary endocrine therapy and those who had surgery plus endocrine therapy, patients who received primary endocrine therapy had worse baseline scores across all domains of the QLQ-C30 (mean(s.d.) global health status score 66.2(21.1) and 77.1(17.8) respectively) (*Table S4*). In the surgery plus endocrine therapy group, a steep drop in mean scores between baseline and 6 weeks was apparent in the role functioning and social functioning domains, and an increase in fatigue and pain scores, compared with values in the primary endocrine therapy group (*Fig. 2a–d*).

Global health status scores (questions 29 and 30 of the instrument), which are calculated from two VAS (ranging from 1 (very poor) to 7 (excellent)) asking the questions 'how would you rate your health today?' and 'how would you rate your quality of life today?' were compared between the unmatched and matched groups. In the unmatched analysis, the observed baseline in global health status in the two groups persisted throughout the 2-year follow-up, although temporal trends owing to treatments were apparent. On matching, the two curves largely overlapped once baseline variation had been accounted for (*Fig. S2*).

Analysis according to whether the surgery was major or minor highlighted the impact of major surgery compared with primary endocrine therapy on several key domains of the QLQ-C30 instrument. The effects of minor or major surgery compared with primary endocrine therapy were small and broadly similar for all four domains across time points (*Fig. 3*). The biggest differences were noted for role functioning; major surgery had a marked negative effect at 6 weeks compared with primary endocrine therapy (mean difference –9.59, 95 per cent c.i. –16.96 to –2.21). This negative effect decreased with time. Fatigue was also more negatively affected by major surgery; levels of fatigue were higher after surgery than those associated with primary endocrine therapy, and took longer to recover. Minor surgery had less impact compared with primary endocrine therapy for both fatigue and role functioning. Levels of pain were higher in the surgery plus endocrine therapy group after both major and minor surgery compared with primary endocrine therapy.

Breast cancer-specific quality-of-life outcomes (EORTC-QLQ-BR23)

Mean(s.d.) scores for each domain of the QLQ-BR23 at each time point are shown in *Table S5*. For most domains, there were no great differences in scores between the groups at 6 weeks, but there were several notable domains where the impact of surgery was apparent (*Fig. 2e,f*). Breast symptom scores in the surgery group increased from a mean(s.d.) of 10.3(13.2) to 23.0(18.8),

Table 1 Patient characteristics of unmatched study population at baseline

	Primary endocrine therapy (*n* = 500)	Surgery (*n* = 2354)	Total (*n* = 2854)
Age (years)	*n* = 500	*n* = 2354	*n* = 2854
Mean (s.d.)	83.5(6.5)	76.4(5.1)	77.6 (6.0)
Median (i.q.r.; range)	84 (79–88; 70–102)	76 (72–80; 69–94)	77 (73–82; 69–102)
aPG-SGA score	*n* = 322	*n* = 2021	*n* = 2343
Mean (s.d.)	2.3(3.1)	1.2(2.2)	1.4(2.4)
Median (i.q.r.; range)	1 (0–3; 0–18)	0 (0–2; 0–17)	0 (0–2; 0–18)
Modified CCI score	*n* = 459	*n* = 2273	*n* = 2732
Mean (s.d.)	5.8(2.0)	4.3(1.4)	4.5(1.6)
Median (i.q.r.; range)	6 (4–7; 3–17)	4 (3–5; 3–13)	4 (3–5; 3–17)
Barthel ADL index score	*n* = 399	*n* = 2135	*n* = 2534
Mean (s.d.)	88.8(16.6)	97.7(6.2)	96.3(9.3)
Median (i.q.r.; range)	95 (85–100; 5–100)	100 (100–100; 10–100)	100.0 (95–100; 5–100)
IADL index score	*n* = 382	*n* = 2104	*n* = 2486
Mean (s.d.)	6.1(2.1)	7.6(0.9)	7.4(1.3)
Median (i.q.r.; range)	7 (5–8; 0–8)	8 (8–8; 0–8)	8 (7–8; 0–8)
MMSE score	*n* = 273	*n* = 1631	*n* = 1904
Mean (s.d.)	26.7(3.7)	28.3(2.5)	28.1(2.8)
Median (i.q.r.; range)	28 (26–29; 10–30)	29 (28–30; 10–30)	29 (27–30; 10–30)
ECOG-PS score	*n* = 453	*n* = 2240	*n* = 2693
Fully active	137 (30.2)	1659 (74.1)	1796 (66.7)
Restricted in physically strenuous activity	167 (36.9)	481 (21.5)	648 (24.1)
Ambulatory and capable of all self-care	67 (14.8)	68 (3.0)	135 (5.0)
Capable of only limited self-care	75 (16.6)	31 (1.4)	106 (3.9)
Completely disabled	7 (1.5)	1 (0.0)	8 (0.3)
No. of current medications	*n* = 450	*n* = 2050	*n* = 2500
Mean (s.d.)	5.0(3.0)	4.1(2.6)	4.2(2.7)
Median (i.q.r.; range)	5 (3–7; 0–18)	4 (2–5; 0–18)	4 (2–6; 0–18)
aPG-SGA risk category	*n* = 376	*n* = 2109	*n* = 2485
Low	291 (77.4)	1848 (87.6)	2139 (86.1)
Moderate	62 (16.5)	227 (10.8)	289 (11.6)
High	23 (6.1)	34 (1.6)	57 (2.3)
ADL risk category	*n* = 408	*n* = 2172	*n* = 2580
No dependency	191 (46.8)	1684 (77.5)	1875 (72.7)
Mild dependency	53 (13.0)	258 (11.9)	311 (12.1)
Moderate/severe dependency	164 (40.2)	230 (10.6)	394 (15.3)
IADL risk category	*n* = 403	*n* = 2158	*n* = 2561
No dependency	158 (39.2)	1759 (81.5)	1917 (74.9)
Mild dependency	64 (15.9)	193 (8.9)	257 (10.0)
Moderate/severe dependency	181 (44.9)	206 (9.5)	387 (15.1)
MMSE risk category	*n* = 500	*n* = 2354	*n* = 2854
Normal function	362 (72.4)	2095 (89.0)	2457 (86.1)
Mild impairment	75 (15.0)	206 (8.8)	281 (9.8)
Moderate impairment	20 (4.0)	29 (1.2)	49 (1.7)
Severe impairment	43 (8.6)	24 (1.0)	67 (2.3)

Values in parentheses are percentages unless indicated otherwise. aPG-SGA, Abridged Patient Generated Subjective Global Assessment; CCI, Charlson Co-morbidity Index; ADL, Activities of Daily Living; IADL, Instrumental ADL; MMSE, Mini-Mental State Examination; ECOG-PS, Eastern Cooperative Oncology Group Performance Status.

whereas a minimal change from 10.1(15.7) to 10.8(15.7) was documented in the primary endocrine therapy group. The score for the surgery plus endocrine therapy group had returned to normal by 24 months.

Patients in the surgery group also had a marked increase in arm symptom score at 6 weeks, whereas those who received primary endocrine therapy did not (increase in mean score from 8.4(14.1) to 16.6(18.3) and from 12.5(18.1) to 14.0(18.9) respectively). Scores in the surgery plus endocrine therapy group had not returned to baseline levels even by 24 months. For both groups, the degree of reduction in body image score was similar when all types of surgery were analysed together. Scores were similar for other domains.

The sexual enjoyment comparison was not valid because of the extremely small number of women who completed these questions, especially in the primary endocrine therapy group.

In the matched cohort, examination of the QLQ-BR23 domains according to minor or major surgery reinforced the impact of surgery on arm function. In longitudinal models fitted to the matched cohort, it was found that women who had major surgery had significantly worse arm symptoms at the 6-week time point than those who had primary endocrine therapy (mean difference 8.85, 95 per cent c.i. 3.63 to 14.07). Women who underwent minor surgery also had more arm symptoms than those who had primary endocrine therapy, but with a smaller difference (*Fig. 4*).

Older age-specific quality-of-life outcomes (EORTC-QLQ-ELD15)

In the unmatched data, patients treated with primary endocrine therapy had worse baseline scores than those who had surgery plus endocrine therapy for mobility problems (mean(s.d.) baseline scores 39.3(33.2) and 15.5(22.5) respectively) (*Table S6*). Similarly, joint stiffness scores were worse in the primary endocrine therapy group at baseline (43.3(35.7) *versus* 27.1(27.9)). Other scores were similar at baseline.

Table 2 Tumour characteristics of unmatched study population at baseline

	Primary endocrine therapy (*n* = 500)	Surgery (*n* = 2354)	Total (*n* = 2854)
No. of clinically involved nodes detectable	*n* = 483	*n* = 2309	*n* = 2792
Mean (s.d.)	0.2(0.6)	0.2(0.8)	0.2(0.7)
Median (i.q.r.; range)	0 (0–0; 0–4)	0 (0–0; 0–20)	0 (0–0; 0–20)
Tumour size (mm)	*n* = 487	*n* = 2318	*n* = 2805
Mean (s.d.)	23.9(12.0)	19.2(12.3)	20.0(12.4)
Median (i.q.r.; range)	21 (16–30; 0–70)	17 (11–24; 0–150)	18 (12–25; 0–150)
Nottingham Prognostic Index score	*n* = 456	*n* = 2172	*n* = 2628
Mean (s.d.)	3.5(0.8)	3.5 (0.8)	3.5 (0.8)
Median (i.q.r.; range)	3.4 (3.2–3.9; 2.1–7.0)	3.3(3.2–4.0; 2.0–6.7)	3.4 (3.2–4.0; 2.0–7.0)
Side of primary tumour	*n* = 500	*n* = 2354	*n* = 2854
Right	223 (44.6)	1084 (46.0)	1307 (45.8)
Left	277 (55.4)	1270 (54.0)	1547 (54.2)
HER2 amplification	*n* = 359	*n* = 1911	*n* = 2270
Negative	311 (86.6)	1641 (85.9)	1952 (86.0)
Inconclusive	14 (3.9)	70 (3.7)	84 (3.7)
Positive	34 (9.5)	200 (10.5)	234 (10.3)
Provisional histological grade	*n* = 484	*n* = 2243	*n* = 2727
1	98 (20.2)	399 (17.8)	497 (18.2)
2	329 (68.0)	1475 (65.8)	1804 (66.2)
3	57 (11.8)	369 (16.5)	426 (15.6)

Values in parentheses are percentages unless indicated otherwise. HER2, human epidermal growth factor receptor 2.

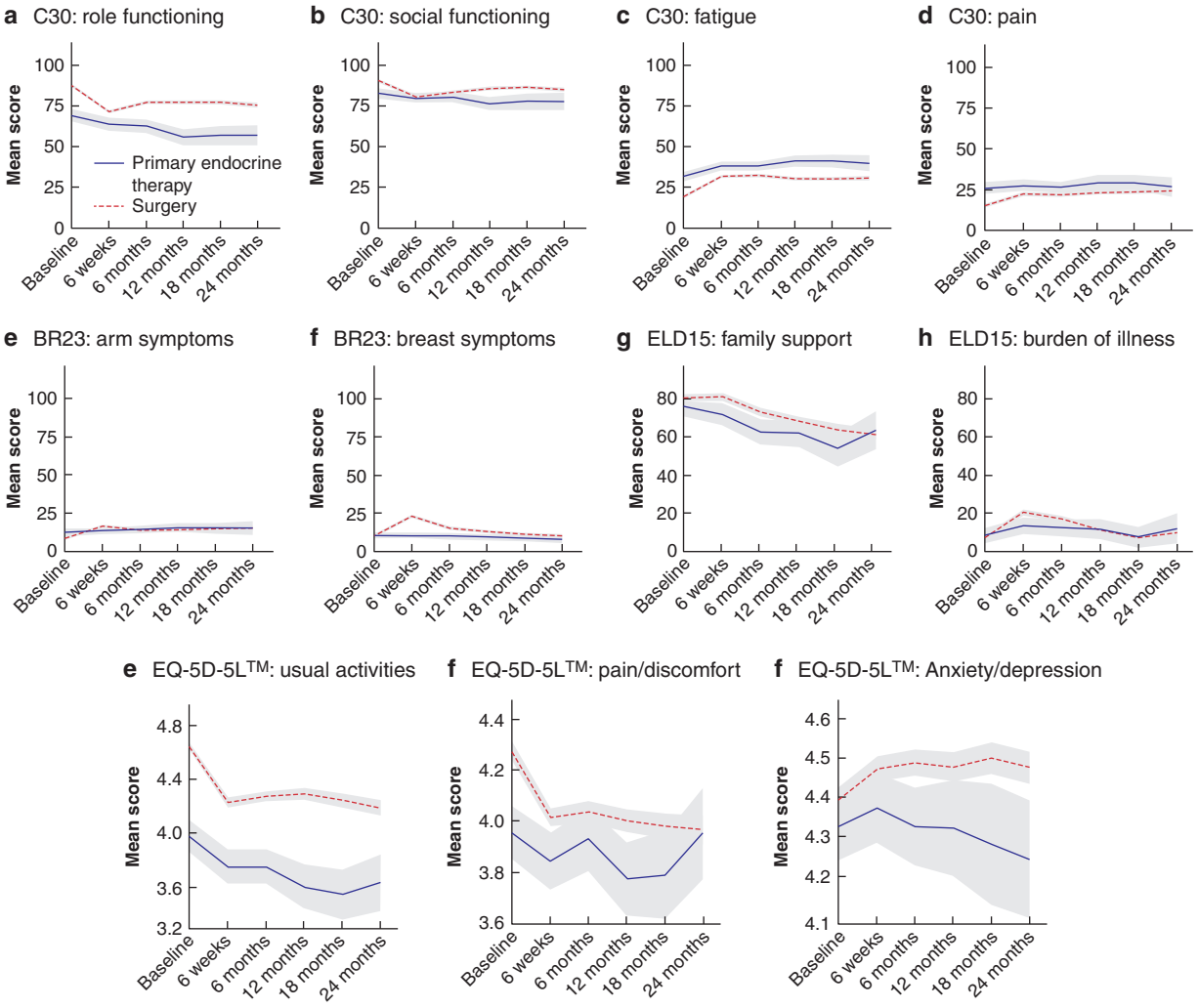

Fig. 2 Unmatched mean quality-of-life scores

QLQ-C30 **a** role functioning, **b** social functioning, **c** fatigue, and **d** pain; QLQ-BR23 **e** arm symptoms and **f** breast symptoms; QLQ-ELD15 **g** family support and **h** burden of illness; EuroQol Five Dimensions 5L (EQ-5D-5L™) **i** usual activities, **j** pain/discomfort, and **k** anxiety/depression. Shaded areas represent 95 per cent confidence intervals.

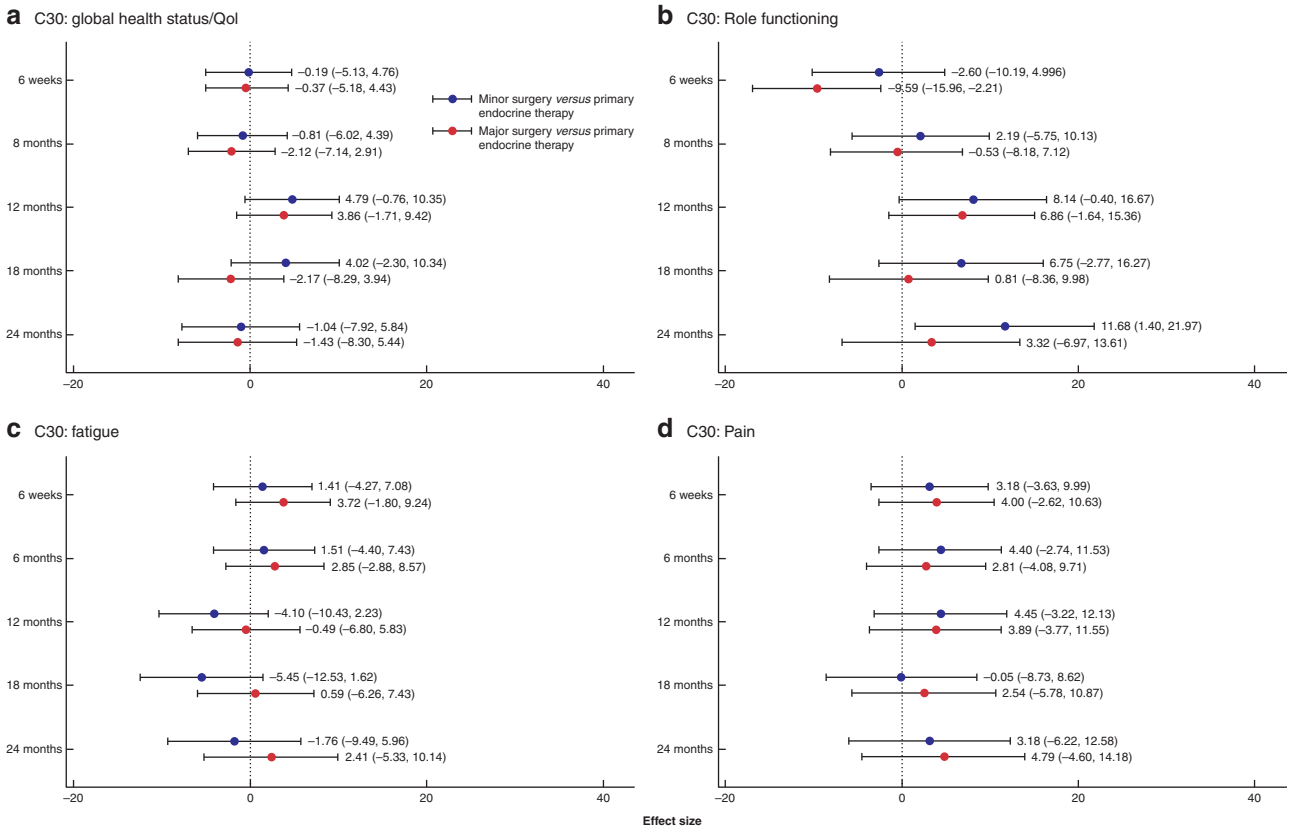

Fig. 3 Longitudinal modelling of effect of minor or major surgery *versus* primary endocrine therapy on QLQ-C30 scores in matched cohort

a Global health status/quality of life (QoL), **c** fatigue, and **d** pain. Effect sizes are shown with 95 per cent confidence intervals.

At the 6-week time point, the most noticeable change in the surgery plus endocrine therapy group was in scores for burden of illness, which increased from a mean of 20.3(23.5) to 30.4(24.7), whereas the score increased from 21.1(23.8) to 25.2(27.5) in the primary endocrine therapy group (*Fig. 2h*). Scores for burden of illness had returned to baseline levels in the surgery group by 18 months.

Domains of interest in the QLQ-ELD15 were analysed using longitudinal modelling in the matched cohort to compare major or minor surgery with primary endocrine therapy. Patients who had major surgery had a greater burden of illness at the 6-week time point than those treated with primary endocrine therapy (mean difference 7.57, 95 per cent c.i. 0.58 to 14.56). In contrast, minor surgery appeared to have a less marked impact (mean difference 5.17, –1.99 to 12.32) (*Fig. 5a*).

EQ-5D-5L™ and visual analogue scale outcomes

Summaries of the overall EQ-5D-5L™ score calculated from all questions and the VAS score showed a substantial difference between baseline scores in the unmatched population owing to imbalances in age and health status (*Table S7*).

Analysis by question showed the impact of surgery on several aspects, notably ability to perform usual activities, pain and discomfort, and anxiety and depression (*Fig. 2i–k*). All of these scores worsened between baseline and 6 weeks when surgery usually took place. It is noteworthy that for none of these domains did the scores return to baseline levels after this treatment had led to deterioration. In the primary endocrine therapy group, the pattern was more of a slow decline over the entire 2-year interval.

In the matched analysis by domain, many of the differences between surgery (major and minor) disappeared (*Fig. 5b–d*).

Discussion

This study included data from a large, multicentre cohort study of women with operable breast cancer treated between 2013 and 2018, 82.5 per cent of whom had surgery as primary treatment. Surgery plus endocrine therapy was shown to have a discernible and clinically valid impact on a range of QoL domains. Women undergoing major surgery had more noticeable changes in pain, role functioning, social functioning, and arm symptoms than those who received primary endocrine therapy which is in keeping with clinical observations. It is noteworthy that, although the majority of scores returned to baseline levels by 24 months, this was not always the case and for some there appeared to be permanent impairment. Again, this accords with clinical observations, particularly in respect of axillary lymph node clearance, after which long-term shoulder stiffness, lymphoedema and pain are relatively common chronic side-effects.

In the primary endocrine therapy group, the pattern of change was largely a slow decline across all domains over the 2-year interval. There was an increase in mobility problems in both groups, which progressed slowly over time, and probably relates to the fact that the majority of the women in both groups were taking an aromatase inhibitor throughout this period, which is known to cause joint pain. This was also seen in the QLQ-ELD15 domain for joint stiffness, where there was a sharp rise in symptoms between 6 weeks and 6 months in both groups; this was in line with most women having started

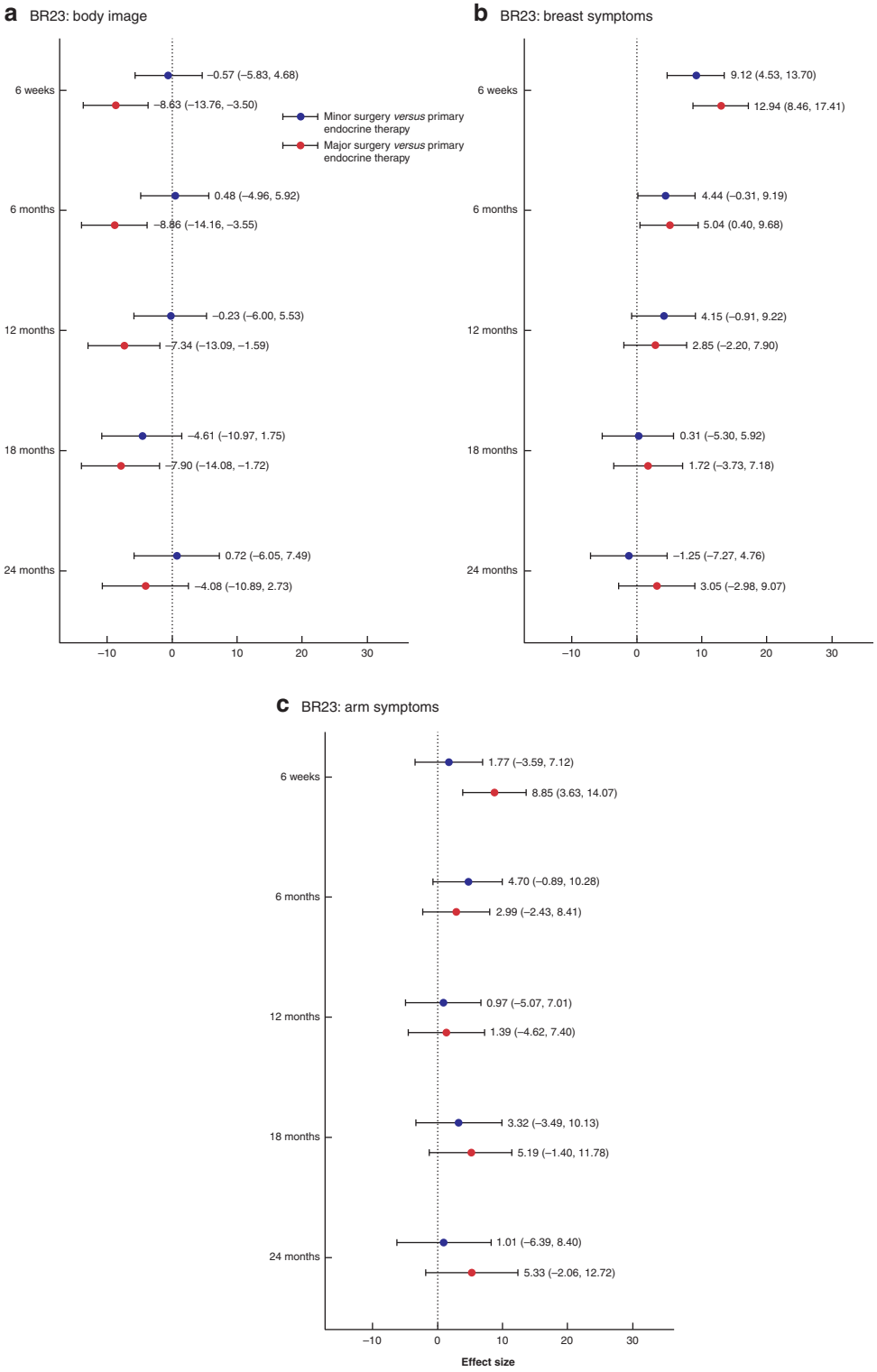

Fig. 4 Longitudinal modelling of effect of minor or major surgery *versus* primary endocrine therapy on QLQ-BR23 scores in matched cohort

a Body image, **c** breast symptoms, and **c** arm symptoms. Effect sizes are shown with 95 per cent confidence intervals.

aromatase inhibitor therapy between baseline and 8 weeks after diagnosis.

The impact of cancer treatment was well reflected in the EQ-5D-5L^{TM} score for usual activities. There was a marked decline between baseline and 6 weeks in the surgery plus endocrine therapy group, but this never recovered and plateaued at this lower level after treatment. In the primary endocrine therapy group, there was just a slow steady decline across the study interval.

For some women, surgery plus endocrine therapy causes both acute and long-term pain, as seen in both the QLQ-C30 and the EQ-5D-5L^{TM} pain domains. Levels of pain increased sharply between baseline and 6 weeks, but never really recovered.

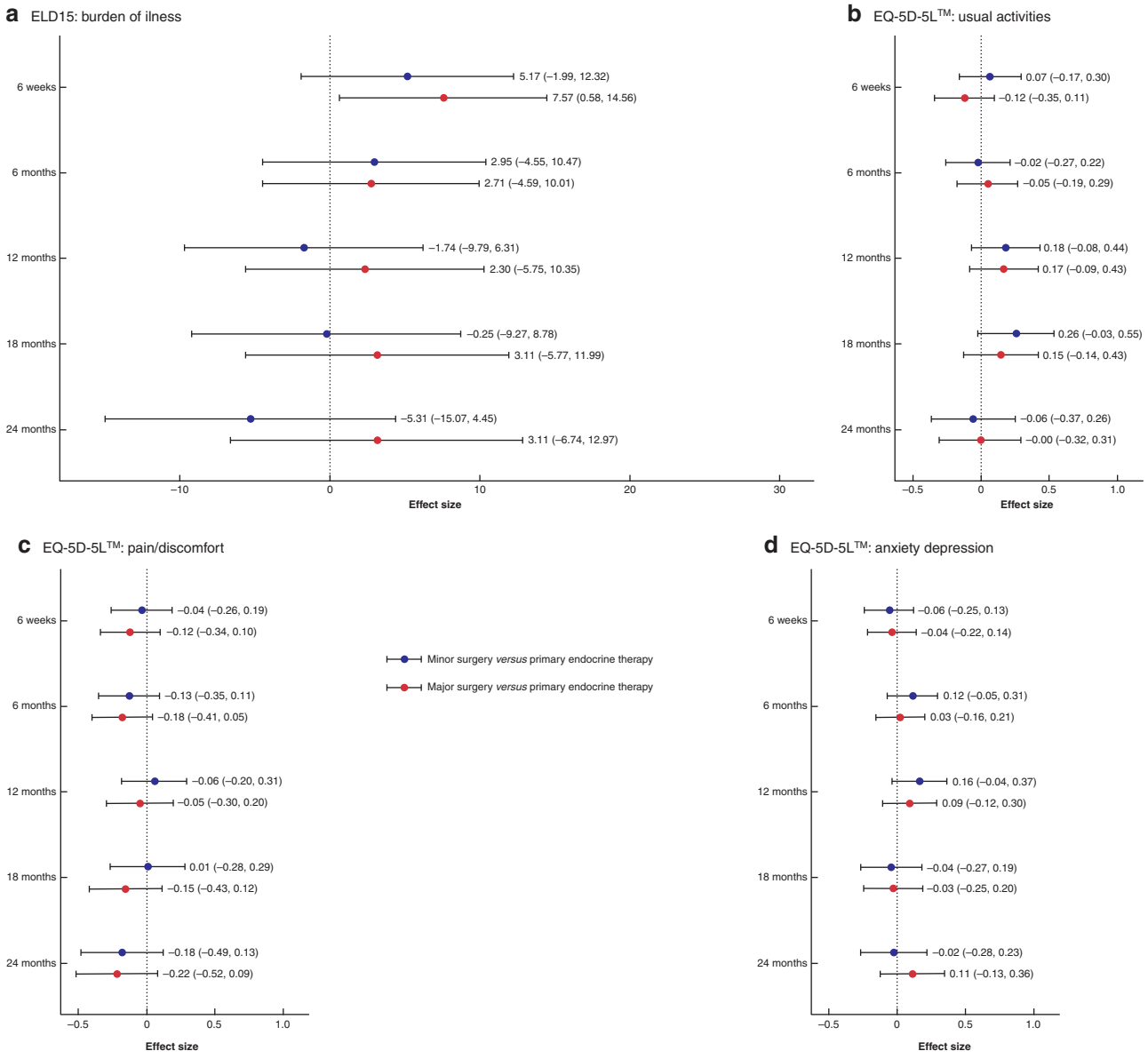

Fig. 5 Longitudinal modelling of effect of minor or major surgery *versus* primary endocrine therapy on QLQ-ELD15 and EuroQol Five Dimensions scores in matched cohort

a ELD15: burden of illness; EuroQol Five Dimensions 5L (EQ-5D-5L™) **b** usual activities, **c** pain/discomfort, and **d** anxiety/depression. Effect sizes are shown with 95 per cent confidence intervals.

The survival data for this study have been published elsewhere[10], and highlight that, for some older frailer patients with breast cancer who have limited life expectancy, standard therapy with surgery plus endocrine therapy may offer limited survival benefit but with a potential reduction in QoL in the short term, with some patients not recovering fully from these effects by 2 years. Women in this borderline group, which appears to include those in moderate or poor health aged over age 85 years, or women older than 90 years (equating to a predicted life expectancy of less than 5 years), should be offered an informed choice of breast cancer treatment and be supported in making the decision based on their own preferences. The Age Gap Decision Tool (https://agegap.shef.ac.uk/) has been developed based on UK registry data; it has been fully validated[27] and tested with older women for usability and acceptability[28–30] to aid in this type of shared decision-making. Recruiting older patients is challenging and many trials fail to recruit well from this age group. The Age Gap study was no exception, and recruited greater numbers of women in their early 70s and fewer in their 90s. Similarly, rates of deaths and withdrawals were higher in the older age groups. In addition, women were able to opt out of QoL form completion to reduce the burden of the trial for frailer patients, and women with severe cognitive impairment were also not expected to complete the forms. Consequently, the study population did not fully represent the UK population of older women and failed to capture QoL outcomes for the oldest and frailest. However, it is based on real-world UK data and represents a large older breast cancer population.

There was attrition in the data collection for QoL forms and different numbers of patients completed the questions at baseline compared with other time points, which could be a confounding factor. The population was heterogeneous with

some older women in excellent health, but others very frail with co-morbidity; consequently, baseline scores in the majority of QoL domains varied widely, with wide standard deviations around mean values because of the heterogeneous population, reduced numbers in the primary endocrine therapy arm of the study and attrition, and also significant differences between treatment groups. This was addressed in part by matching, but this process reduced the number of patients for analysis and therefore the power of the study. The matching process provided information on a group for whom UK surgeons seem to have collective equipoise about optimal treatment selection, but also excluded very fit surgical patients and the least fit patients treated with primary endocrine therapy, for whom these comparisons may not be valid. The ideal way of studying this issue would be to undertake an RCT. Such a trial (ESTEEM) was attempted, but was unsuccessful owing to poor recruitment as a consequence of lack of patient and clinician equipoise[31].

A further complexity was the variation in treatments in the surgery plus endocrine therapy group. These patients had surgery varying between mastectomy and reconstruction plus axillary lymph node clearance at one extreme and wide local excision only at the other. Some women also had chemotherapy and radiotherapy. 'Surgery plus endocrine therapy' should be more appropriately referred to as 'standard therapy'. An attempt was made to address this by subcategorization of surgery as major or minor, with a focus on the 6-week time point, at which stage the majority of women had only undergone surgery and not adjuvant radiotherapy or chemotherapy. Detailed analysis of the impact of chemotherapy and radiotherapy will be the subject of separate analyses. The temporal trends showing the impact of surgery were clinically as expected and therefore it was felt that the conclusions are valid.

From this large study of the impact of treatment on QoL in older women with operable breast cancer, it is clear that breast cancer treatment has a negative impact on QoL, which is most notable after surgery. For most patients, QoL gradually returns to baseline level, but some residual impairment remains at 24 months after diagnosis. It is known that surgery is more effective than primary endocrine therapy in curing and controlling breast cancer in most women, with a significant benefit in terms of local control and survival. This must be traded against a reduction in short-term QoL, and women must be made aware that they may suffer some longer-term impairments. There is evidence that QoL and preservation of independence are of relatively higher value for older women than for younger women[3]. Frailer women with co-morbidities and a reduced life expectancy may gain little survival benefit from surgery, and potentially suffer long-term deterioration in QoL and function as a result of the operation, which should be taken into account when supporting shared decision-making.

Acknowledgements

J.L.M. and A.S. are joint first authors of this article. The trial sponsor was Doncaster and Bassetlaw Teaching Hospitals NHS Foundation Trust (Doncaster Royal Infirmary, Doncaster, UK). This study was funded by the UK National Institute for Health Research (NIHR). This paper presents independent research funded by the NIHR under its Programme Grants for Applied Research Programme (grant reference number RP-PG-1209-10071). T.R. and S.J.W. are NIHR Senior Investigators, and J.L.M. is a NIHR Clinical Lecturer. The views expressed are those of the authors and not necessarily those of the NIHR or the Department of Health and Social Care.

Disclosure: The authors declare no conflict of interest.

Supplementary material

Supplementary material is available at *BJS* online.

References

1. Preece PE, Wood RA, Mackie CR, Cuschieri A. Tamoxifen as initial sole treatment of localised breast cancer in elderly women: a pilot study. *Br Med J (Clin Res Ed)* 1982;**284**:869–870
2. Derks MGM, Bastiaannet E, Kiderlen M, Hilling DE, Boelens PG, Walsh PM *et al.* Variation in treatment and survival of older patients with non-metastatic breast cancer in five European countries: a population-based cohort study from the EURECCA Breast Cancer Group. *Br J Cancer* 2018; **119**:121–129
3. Husain LS, Collins K, Reed M, Wyld L. Choices in cancer treatment: a qualitative study of the older women's (>70 years) perspective. *Psychooncology* 2008;**17**:410–416
4. Biganzoli L, Wildiers H, Oakman C, Marotti L, Loibl S, Kunkler I *et al.* Management of elderly patients with breast cancer: updated recommendations of the International Society of Geriatric Oncology (SIOG) and European Society of Breast Cancer Specialists (EUSOMA). *Lancet Oncol* 2012;**13**:e148–e160
5. Fallowfield L. Quality of life in the elderly woman with breast cancer treated with tamoxifen and surgery or tamoxifen alone. *Journal of Womens Health* 1994;**3**:17–20
6. Aaronson NK, Ahmedzai S, Bergman B, Bullinger M, Cull A, Duez NJ *et al.* The European Organization for Research and Treatment of Cancer QLQ-C30: a quality-of-life instrument for use in international clinical trials in oncology. *J Natl Cancer Inst* 1993;**85**:365–376
7. Sprangers MA, Groenvold M, Arraras JI, te Velde A, Muller M, Franzini L *et al.* The European Organization for Research and Treatment of Cancer breast cancer-specific quality-of-life questionnaire module: first results from a three-country field study. *J Clin Oncol* 1996;**14**:2756–2768
8. Johnson C, Fitzsimmons D, Gilbert J, Arrarras JI, Hammerlid E, Bredart A *et al.* Development of the European Organisation for Research and Treatment of Cancer quality of life questionnaire module for older people with cancer: the EORTC QLQ-ELD15. *Eur J Cancer* 2010;**46**:2242–2252
9. NHS Information Centre. *National Mastectomy and Breast Reconstruction Audit Fourth Annual Report; Royal College of Surgeons of England 2011*. https://files.digital.nhs.uk/publicationimport/pub02xxx/pub02731/clin-audi-supp-prog-mast-brea-reco-2011-rep1.pdf
10. Morgan JL, George J, Holmes G, Martin C, Reed MWR, Ward S *et al.* Breast cancer surgery in older women: outcomes of the bridging age gap in breast cancer study. *Br J Surg* 2020;**107**:1468–1479
11. Tang V, Zhao S, Boscardin J, Sudore R, Covinsky K, Walter LC *et al.* Functional status and survival after breast cancer surgery in nursing home residents. *JAMA Surg* 2018;**153**:1090–1096
12. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP *et al.* The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *Lancet* 2007;**370**:1453–1457

13. Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *J Chronic Dis* 1987;**40**:373–383
14. Read JA, Crockett N, Volker DH, MacLennon P, Boris Choy ST, Beale P *et al.* Nutritional assessment in cancer: comparing the Mini-Nutritional Assessment (MNA) with the scored Patient-Generated Subjective Global Assessment (PGSGA). *Nutr Cancer* 2005;**53**:51–56
15. Ottery FD. Definition of standardized nutritional assessment and interventional pathways in oncology. *Nutrition* 1996;**12**: S15–S19
16. Vigano AL, di Tomasso J, Kilgour RD, Trutschnigg B, Lucar E, Morais JA *et al.* The abridged patient-generated subjective global assessment is a useful tool for early detection and characterization of cancer cachexia. *J Acad Nutr Diet* 2014 **114**:1088–1098
17. Mahoney FI, Barthel DW. Functional evaluation: the Barthel Index. *Md State Med J* 1965;**14**:61–65
18. Lawton MP, Brody EM. Assessment of older people: self-maintaining and instrumental activities of daily living. *Gerontologist* 1969;**9**:179–186
19. Folstein MF, Folstein SE, McHugh PR. 'Mini-mental state'. A practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res* 1975;**12**:189–198
20. Oken MM, Creech RH, Tormey DC, McFadden ET, Carbone PP. Toxicity and response criteria of the Eastern Cooperative Oncology Group. *Am J Clin Oncol* 1982;**5**:649–655
21. Herdman M, Gudex C, Lloyd A, Janssen M, Kind P, Parkin D, Bonsel G *et al.* Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). *Qual Life Res* 2011; **20**: 1727–1736
22. Therasse P, Arbuck SG, Eisenhauer EA, Wanders J, Kaplan RS, Rubinstein L *et al.* New guidelines to evaluate the response to treatment in solid tumors. European Organization for Research and Treatment of Cancer, National Cancer Institute of the United States, National Cancer Institute of Canada. *J Natl Cancer Inst* 2000;**92**:205–216
23. Fayers PM, Aaronson NK, Bjordal K, Groenvold M, Curran D, Bottomley A. *The EORTC QLQ-C30 Scoring Manual* (3rd edn). Brussels: European Organisation for Research and Treatment of Cancer, 2001.
24. Wheelwright S, Darlington AS, Fitzsimmons D, Fayers P, Arraras JI, Bonnetain F *et al.* International validation of the EORTC QLQ-ELD14 questionnaire for assessment of health-related quality of life elderly patients with cancer. *Br J Cancer* 2013;**109**:852–858
25. Devlin NJ, Shah KK, Feng Y, Mulhern B, van Hout B. Valuing health-related quality of life: an EQ-5D-5L value set for England. *Health Econ* 2018;**27**:7–22
26. Galea MH, Blamey RW, Elston CE, Ellis IO. The nottingham prognostic index in primary breast cancer. Breast Cancer Res Treat 1992;**22**:207–19
27. Ward SE, Richards PD, Morgan JL, Holmes GR, Broggio JW, Collins K *et al.* Omission of surgery in older women with early breast cancer has an adverse impact on breast cancer-specific survival. *Br J Surg* 2018;**105**:1454–1463
28. Burton M, Collins KA, Lifford KJ, Brain K, Wyld L, Caldon L *et al.* The information and decision support needs of older women (>75 yrs) facing treatment choices for breast cancer: a qualitative study. *Psychooncology* 2015;**24**:878–884
29. Burton M, Kilner K, Wyld L, Lifford KJ, Gordon F, Allison A *et al.* Information needs and decision-making preferences of older women offered a choice between surgery and primary endocrine therapy for early breast cancer. *Psychooncology* 2017;**26**: 2094–2100
30. Lifford KJ, Edwards A, Burton M, Harder HG, Armitage F, Morgan JL *et al.* Efficient development and usability testing of decision support interventions for older women with breast cancer. *Patient Prefer Adherence* 2019;**13**:131–143
31. Reed MW, Wyld L, Ellis P, Bliss J, Leonard R. Breast cancer in older women: trials and tribulations. *Clin Oncol (R Coll Radiol)* 2009;**2**:99–102